

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 9/00</b>		A2	(11) International Publication Number: <b>WO 99/53038</b> (43) International Publication Date: 21 October 1999 (21.10.99)
(21) International Application Number: PCT/US99/08253		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 14 April 1999 (14.04.99)		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(30) Priority Data: 09/060,872 15 April 1998 (15.04.98) US			
(71) Applicant: GENENCOR INTERNATIONAL, INC. [US/US]; 4 Cambridge Place, 1870 South Winton Road, Rochester, NY 14618 (US).			
(72) Inventors: ESTELL, David, A.; 248 Woodbridge Circle, San Mateo, CA 94403 (US). HARDING, Fiona, A.; 772 Lewis Street, Santa Clara, CA 95050 (US).			
(74) Agent: STONE, Christopher, L.; Genencor International, Inc., 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).			
(54) Title: MUTANT PROTEINS HAVING LOWER ALLERGENIC RESPONSE IN HUMANS AND METHODS FOR CONSTRUCTING, IDENTIFYING AND PRODUCING SUCH PROTEINS			
(57) Abstract <p>The present invention relates to a novel improved protein mutant which produces low allergenic response in humans compared to the parent of that mutant. Specifically, the present invention comprises neutralizing or reducing the ability of T-cells to recognize epitopes and thus prevent sensitization of an individual to the protein.</p> <p><i>See p. 29 all rights reserved in process 2000 PCT published</i></p>			

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Republic of Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**MUTANT PROTEINS HAVING LOWER ALLERGENIC  
RESPONSE IN HUMANS AND METHODS FOR  
CONSTRUCTING, IDENTIFYING AND PRODUCING SUCH PROTEINS**

5

**BACKGROUND OF THE INVENTION**

**A. Field of the Invention**

The present invention relates to proteins which produce lower allergenic response in humans exposed to such proteins, and an assay predictive of such response. More specifically, the present invention relates to a novel improved protein mutant which 10 produces very low allergenic response in humans sensitized to that protein through exposure compared to the precursor of such protein mutant.

**B. State of the Art**

Proteins used in industrial, pharmaceutical and commercial applications are of 15 increasing prevalence. As a result, the increased exposure due to this prevalence has been responsible for some safety hazards caused by the sensitization of certain persons to those peptides, whereupon subsequent exposure causes extreme allergic reactions which can be injurious and even fatal. For example, proteases are known to cause dangerous hypersensitivity in some individuals. As a result, despite the usefulness of 20 proteases in industry, e.g., in laundry detergents, cosmetics, textile treatment etc...., and the extensive research performed in the field to provide improved proteases which have, for example, more effective stain removal under detergency conditions, the use of proteases in industry has been problematic due to their ability to produce a hypersensitive allergenic response in some humans.

25        Much work has been done to alleviate these problems. Among the strategies explored to reduce immunogenic potential of protease use have been improved production processes which reduce potential contact by controlling and minimizing workplace concentrations of dust particles or aerosol carrying airborne protease, improved granulation processes which reduce the amount of dust or aerosol actually produced from 30 the protease product, and improved recovery processes to reduce the level of potentially allergenic contaminants in the final product. However, efforts to reduce the allergenicity of protease, per se, have been relatively unsuccessful. Alternatively, efforts have been made to mask epitopes in protease which are recognized by immunoglobulin E (IgE) in hypersensitive individuals (PCT Publication No. WO 92/10755) or to enlarge or change the 35 nature of the antigenic determinants by attaching polymers or peptides/proteins to the problematic protease.

When an adaptive immune response occurs in an exaggerated or inappropriate form, the individual experiencing the reaction is said to be hypersensitive. Hypersensitivity reactions are the result of normally beneficial immune responses acting inappropriately and sometimes cause inflammatory reactions and tissue damage. They can be provoked 5 by many antigens; and the cause of a hypersensitivity reaction will vary from one individual to the next. Hypersensitivity does not normally manifest itself upon first contact with the antigen, but usually appears upon subsequent contact. One form of hypersensitivity occurs when an IgE response is directed against innocuous environmental antigens, such as pollen, dust-mites or animal dander. The resulting release of 10 pharmacological mediators by IgE-sensitized mast cells produces an acute inflammatory reaction with symptoms such as asthma or rhinitis.

Nonetheless, a strategy comprising modifying the IgE sites will not generally be successful in preventing the cause of the initial sensitization reaction. Accordingly, such strategies, while perhaps neutralizing or reducing the severity of the subsequent 15 hypersensitivity reaction, will not reduce the number or persons actually sensitized. For example, when a person is known to be hypersensitive to a certain antigen, the general, and only safe, manner of dealing with such a situation is to isolate the hypersensitive person from the antigen as completely as possible. Indeed, any other course of action would be dangerous to the health of the hypersensitive individual. Thus, while reducing 20 the danger of a specific protein for a hypersensitive individual is important, for industrial purposes it would be far more valuable to render a protein incapable of initiating the hypersensitivity reaction in the first place.

T-lymphocytes (T-cells) are key players in the induction and regulation of immune responses and in the execution of immunological effector functions. Specific immunity 25 against infectious agents and tumors is known to be dependent on these cells and they are believed to contribute to the healing of injuries. On the other hand, failure to control these responses can lead to auto aggression. In general, antigen is presented to T-cells in the form of antigen presenting cells which, through a variety of cell surface mechanisms, capture and display antigen or partial antigen in a manner suitable for 30 antigen recognition by the T-cell. Upon recognition of a specific epitope by the receptors on the surface of the T-cells (T-cell receptors), the T-cells begin a series of complex interactions, including proliferation, which result in the production of antibody by B-cells. While T-cells and B-cells are both activated by antigenic epitopes which exist on a given protein or peptide, the actual epitopes recognized by these mononuclear cells are 35 generally not identical. In fact, the epitope which activates a T-cell to initiate the creation

of immunologic diversity is quite often not the same epitope which is later recognized by B-cells in the course of the immunologic response. Thus, with respect to hypersensitivity, while the specific antigenic interaction between the T-cell and the antigen is a critical element in the initiation of the immune response to antigenic exposure, the specifics of 5 that interaction, i.e., the epitope recognized, is often not relevant to subsequent development of a full blown allergic reaction.

PCT Publication No. WO 96/40791 discloses a process for producing polyalkylene oxide-polypeptide conjugates with reduced allergenicity using polyalkylene oxide as a starting material.

10 PCT Publication No. WO 97/30148 discloses a polypeptide conjugate with reduced allergenicity which comprises one polymeric carrier molecule having two or more polypeptide molecules coupled covalently thereto.

15 PCT Publication No. WO 96/17929 discloses a process for producing polypeptides with reduced allergenicity comprising the step of conjugating from 1 to 30 polymolecules to a parent polypeptide.

PCT Publication No. WO 92/10755 discloses a method of producing protein variants evoking a reduced immunogenic response in animals. In this application, the proteins of interest, a series of proteases and variants thereof, were used to immunized rats. The sera from the rats was then used to measure the reactivity of the polyclonal 20 antibodies already produced and present in the immunized sera to the protein of interest and variants thereof. From these results, it was possible to determine whether the antibodies in the preparation were comparatively more or less reactive with the protein and its variants, thus permitting an analysis of which changes in the protein are likely to neutralize or reduce the ability of the Ig to bind. From these tests on rats, the conclusion 25 was arrived at that changing any of subtilisin 309 residues corresponding to 127, 128, 129, 130, 131, 151, 136, 151, 152, 153, 154, 161, 162, 163, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 186, 193, 194, 195, 196, 197, 247, 251, 261 will result in a change in the immunological potential.

30 PCT Publication No. WO 94/10191 discloses low allergenic proteins comprising oligomeric forms of the parent monomeric protein, wherein the oligomer has substantially retained its activity.

The prior art has provided methods of reducing the allergenicity of certain proteins and identification of epitopes which cause allergic reactions in some individuals, the assays used to identify these epitopes generally involving measurement of IgE and IgG 35 antibody in blood sera previously exposed to the antigen. Nonetheless, once an Ig

reaction has been initiated, sensitization has already occurred. Accordingly, there is a need for a method of determining epitopes which cause sensitization in the first place, as neutralization of these epitopes will result in significantly less possibility for sensitization to occur, thus reducing the possibility of initial sensitization.

5

### SUMMARY OF THE INVENTION

It is an object of the invention to provide a protein having decreased potential to cause allergenic response in humans compared to a precursor protein.

It is a further object of the present invention to provide for a protease variant which 10 has useful activity in common protease applications, such as detergents and or the treatment of wool to prevent felting, in bar or liquid soap applications, dish-care formulations, contact lens cleaning solutions or products, peptide hydrolysis, waste treatment, textile applications such as anti-felting, in cosmetic formulations and for skin care, or as fusion-cleavage enzymes in protein production, which protease variant can be 15 more safely used due to its lowered allergenic potential.

According to the present invention, a method for identifying T-cell epitopes within a protein is provided. The present invention provides an assay which identifies epitopes as follows: antigen presenting cells are combined with naïve human T-cells and with a peptide of interest. In a preferred embodiment of the invention, a method is provided 20 wherein a T-cell epitope is recognized comprising the steps of: (a) obtaining from a single blood source a solution of dendritic cells and a solution of naïve CD4+ and/or CD8+ T-cells; (b) promoting differentiation in said solution of dendritic cells; (c) combining said solution of differentiated dendritic cells and said naïve CD4+ and/or CD8+ T-cells with a peptide of interest; (d) measuring the proliferation of T-cells in said step (c).

25 According to another embodiment of the present invention, a protein is provided in which a T-cell epitope is modified so as to reduce or preferably neutralize (eliminate) the ability of the T-cell to identify that epitope. Thus, a protein is provided having reduced allergenicity, wherein said protein comprises a modification comprising the substitution or deletion of amino acid residues which are identified as within a T-cell epitope. According 30 to a preferred embodiment, an epitope is determined in a protein or peptide which, when recognized by a T-cell, results in the proliferation of T-cells which is greater than the baseline. That T-cell epitope is then modified so that, when the peptide comprising the epitope is analyzed in the assay of the invention, it results in lesser proliferation than the protein comprising the unmodified epitope. More preferably, the epitope to be modified 35 produces greater than three times the baseline T-cell proliferation in a sample. When

modified, the epitope produces less than three times the baseline T-cell proliferation, preferably less than two times the baseline T-cell proliferation and most preferably less than or substantially equal to the baseline T-cell proliferation in a sample.

Preferably, the epitope is modified in one of the following ways: (a) the amino acid sequence of the epitope is substituted with an analogous sequence from a human homolog to the protein of interest, i.e., human subtilisin or another human protease derived subtilisin like molecule such as furin or the kexins (see e.g., *Methods in Enzymology*, Vol. 244., (1994) pp. 175 et seq; Roebroek et al., *EMBO J.*, Vol. 5, No. 9, pp. 2197-2202 (1986); Tomkinson et al., *Biochem.*, Vol. 30, pp. 168-174 (1991); Keifer et al., *DNA and Cell Biol.*, Vol. 10, No. 10, pp. 757-769 (1991)); (b) the amino acid sequence of the epitope is substituted with an analogous sequence from a non-human homolog to the protein of interest, which analogous sequence produces a lesser allergenic response due to T-cell recognition than that of the protein of interest; (c) the amino acid sequence of the epitope is substituted with a sequence which substantially mimics the major tertiary structure attributes of the epitope, but which produces a lesser allergenic response due to T-cell recognition than that of the protein of interest; or (d) with any sequence which produces lesser allergenic response due to T-cell recognition than that of the protein of interest.

In a specific embodiment of the invention, a protease variant is provided comprising at least one amino acid substitution at a position corresponding to residues 170, 171, 172 and/or 173 in BPN', wherein such substitutions comprise modifying residue 170 to aspartic acid, modifying residue 171 to glutamine, modifying residue 172 to methionine and/or modifying residue 173 to aspartic acid. In a most preferred embodiment, the substitution comprises modifying residues 170, 171 and 173 to aspartic acid, glutamine and aspartic acid, respectively.

In another embodiment of the present invention, a method for producing the protein of the invention having reduced allergenicity is provided. Preferably, the mutant protein is prepared by modifying a DNA encoding a precursor protein so that the modified DNA encodes the mutant protein of the invention.

In yet another embodiment of the invention, DNA sequences encoding the mutant protein, as well as expression vectors containing such DNA sequences and host cells transformed with such vectors are provided, which host cells are preferably capable of expressing such DNA to produce the mutant protein of the invention either intracellularly or extracellularly.

The mutant protein of the invention is useful in any composition or process in which the precursor protein is generally known to be useful. For example, where the protein is a protease, the reduced allergenicity protease can be used as a component in cleaning products such as laundry detergents and hard surface cleansers, as an aid in the preparation of leather, in the treatment of textiles such as wool and/or silk to reduce felting, as a component in a personal care, cosmetic or face cream product, and as a component in animal or pet feed to improve the nutritional value of the feed. Similarly, where the protein is an amylase, the reduced allergenicity amylase can be used for the liquefaction of starch, as a component in a dishwashing detergent, for desizing of textiles, in a laundry detergent or any other use for which amylase is useful.

One advantage of the present invention is that by measuring the proliferation of T-cells due to T-cell epitope recognition, it is possible to identify peptides which contain epitopes responsible for initially sensitizing an individual. That is, T-cell proliferation due to T-cell epitope recognition results in sensitization of an individual to that peptide or a protein which contains it. Neutralization of such "sensitizing" T-cell epitopes will inevitably result in a greater degree of safety for those who handle or are otherwise exposed to the antigen containing the epitope because they will not be initially sensitized, thus preventing the production of Ig antibodies typical of an allergic reaction upon subsequent exposure to the antigen.

An advantage of the present invention is the preparation of proteins, including enzymes, which may be used with significantly less danger of sensitization for the individuals exposed. Thus, for example, the proteins of the invention may be more safely used in cosmetics such as face creams, detergents such as laundry detergents, hard surface cleaning compositions and pre-wash compositions or any other use of protein, including enzymes, wherein human exposure is a necessary by-product.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1 A, B1, B2 and B3 illustrate the DNA (SEQ ID: NO 1) and amino acid (SEQ ID: NO 2) sequence for *Bacillus amyloliquefaciens* subtilisin (BPN') and a partial restriction map of this gene.

Fig. 2 illustrates the conserved amino acid residues among subtilisins from *Bacillus amyloliquefaciens* (SEQ ID: NO 3) and *Bacillus lentus* (wild-type) (SEQ ID: NO 4).

Figs. 3A and 3B illustrate an amino acid sequence alignment of subtilisin type proteases from *Bacillus amyloliquefaciens* (BPN'), *Bacillus subtilis*, *Bacillus licheniformis*

(SEQ ID:NO 5) and *Bacillus lenthus*. The symbol \* denotes the absence of specific amino acid residues as compared to subtilisin BPN'.

Fig. 4 illustrates the additive T-cell response of 16 peripheral mononuclear blood samples to peptides corresponding to the *Bacillus lenthus* protease. Peptide E05 includes the region comprising residues corresponding to 170-173 in protease from *Bacillus amyloliquefaciens*.

Fig. 5 illustrates the additive T-cell response of 10 peripheral mononuclear blood samples to peptides corresponding to the human subtilisin molecule. Peptides F10, F9, F8 and F7 all contain the amino acid sequence DQMD corresponding to the region comprising residues corresponding to 170-173 in protease from *Bacillus amyloliquefaciens* in the sequence alignment of Fig. 3.

Fig. 6A and 6B/6C illustrate amino acid strings corresponding to peptides derived from the sequence of *Bacillus lenthus* protease and a human subtilisin, respectively.

Fig. 7 illustrates the amino acid sequence of human subtilisin (SEQ ID:NO 6).

Fig. 8 illustrates an amino acid sequence alignment of BPN' (*Bacillus amyloliquefaciens*) protease, SAVINASE (*Bacillus lenthus*) protease and human subtilisin (S2HSBT).

Fig. 9 illustrates the T-cell response to peptides derived from *Bacillus lenthus* protease in a sample taken from an individual known to be hypersensitive to *Bacillus lenthus* protease. Peptide E05 represents the region corresponding to 170-173 in protease from *Bacillus amyloliquefaciens*.

Fig. 10 illustrates the T-cell response to various alanine substitutions in the E05 *Bacillus lenthus* protease peptide set in a sample taken from an individual known to be hypersensitive to *Bacillus lenthus* protease.

#### DETAILED DESCRIPTION OF THE INVENTION

According to the present invention, a method for identifying T-cell epitopes is provided. The present invention provides an assay which identifies epitopes as follows: differentiated dendritic cells are combined with naïve human CD4+ and/or CD8+ T-cells and with a peptide of interest. More specifically, a method is provided wherein a T-cell epitope is recognized comprising the steps of: (a) obtaining from a single blood source a solution of dendritic cells and a solution of naïve CD4+ and/or CD8+ T-cells; (b) promoting differentiation in said solution of dendritic cells; (c) combining said solution of differentiated

dendritic cells and said naïve CD4+ and/or CD8+ T-cells with a peptide of interest; (d) measuring the proliferation of T-cells in said step (c).

The peptide of interest to be analyzed according to the assay of the invention is derived from a protein or enzyme for which reduced allergenicity is desirable or required.

- 5 In the practice of the invention, it is possible to identify with precision the location of an epitope which can cause sensitization in an individual or sampling of individuals. In a particularly effective embodiment of the invention, a series of peptide oligomers which correspond to all or part of the protein or enzyme are prepared. For example, a peptide library is produced covering the relevant portion or all of the protein. One particularly useful manner of producing the peptides is to introduce overlap into the peptide library, for example, producing a first peptide corresponds to amino acid sequence 1-10 of the subject protein, a second peptide corresponds to amino acid sequence 4-14 of the subject protein, a third peptide corresponds to amino acid sequence 7-17 of the subject protein, a fourth peptide corresponds to amino acid sequence 10-20 of the subject protein etc....until 10 representative peptides corresponding to the entire molecule are created. By analyzing each of the peptides individually in the assay provided herein, it is possible to precisely identify the location of epitopes recognized by T-cells. In the example above, the reaction of one specific peptide to a greater extent than its neighbors will facilitate identification of the epitope anchor region to within three amino acids. After determining the location of 15 these epitopes, it is possible to alter the amino acids within each epitope until the peptide produces a less significant T-cell response.
- 20
- 25

"Antigen presenting cell" as used herein means a cell of the immune system which present antigen on their surface which is recognizable by receptors on the surface of T-cells. Examples of antigen presenting cells are dendritic cells, interdigitating cells, activated B-cells and macrophages.

"T-cell proliferation" as used herein means the number of T-cells produced during the incubation of T-cells with the antigen presenting cells, with or without antigen.

- 30 "Baseline T-cell proliferation" as used herein means T-cell proliferation which is normally seen in an individual in response to exposure to antigen presenting cells in the absence of peptide or protein antigen. For the purposes herein, the baseline T-cell proliferation level was determined on a per sample basis for each individual as the proliferation of T-cells in response to antigen presenting cells in the absence of antigen.
- 35

"T-cell epitope" means a feature of a peptide or protein which is recognized by a T-cell receptor in the initiation of an immunologic response to the peptide comprising that antigen. Recognition of a T-cell epitope by a T-cell is generally believed to be via a

mechanism wherein T-cells recognize peptide fragments of antigens which are bound to class I or class II major histocompatibility (MHC) molecules expressed on antigen-presenting cells (see e.g., Moeller, G. ed., *Antigenic Requirements for Activation of MHC-Restricted Responses, Immunological Review*, Vol. 98, p. 187 (Copenhagen; 5 Munksgaard) (1987).

The epitopes determined according to the assay provided herein are then modified to reduce the allergenic potential of the protein of interest. In a preferred embodiment, the epitope to be modified produces a level of T-cell proliferation of greater than three times the baseline T-cell proliferation in a sample. When modified, the epitope produces less 10 than three times the baseline proliferation, preferably less than two times the baseline proliferation and most preferably less than or substantially equal to the baseline proliferation in a sample.

Preferably, the epitope is modified in one of the following ways: (a) the amino acid sequence of the epitope is substituted with an analogous sequence from a human 15 homolog to the protein of interest; (b) the amino acid sequence of the epitope is substituted with an analogous sequence from a non-human homolog to the protein of interest, which analogous sequence produces a lesser allergenic response due to T-cell epitope recognition than that of the protein of interest; (c) the amino acid sequence of the epitope is substituted with a sequence which substantially mimics the major tertiary 20 structure attributes of the epitope, but which produces a lesser allergenic response due to T-cell epitope recognition than that of the protein of interest; or (d) with any sequence which produces lesser allergenic response due to T-cell epitope recognition than that of the protein of interest.

“Sample” as used herein comprises mononuclear cells which are naïve, i.e., not 25 sensitized, to the antigen in question.

“Homolog” as used herein means a protein or enzyme which has similar catalytic action, structure and/or use as the protein of interest. It is desirable to find a homolog that has a tertiary and/or primary structure similar to the protein of interest as replacement of the epitope in the protein of interest with an analogous segment from the homolog will 30 reduce the disruptiveness of the change. Thus, closely homologous enzymes will provide the most desirable source of epitope substitutions. Alternatively, if possible, it is advantageous to look to human analogs for a given protein. For example, substituting a specific epitope in a bacterial subtilisin with a sequence from a human analog to subtilisin (i.e., human subtilisin) should result in less allergenicity in the bacterial protein.

An "analogous" sequence may be determined by ensuring that the replacement amino acids show a similar function, the tertiary structure and/or conserved residues to the amino acids in the protein of interest at or near the epitope. Thus, where the epitope region contains, for example, an alpha-helix or a beta-sheet structure, the replacement 5 amino acids should maintain that specific structure.

While the present invention extends to all proteins for which it is desired to reduce allergenicity, for the sake of simplicity, the following will describe a particularly preferred embodiment of the invention, the modification of protease. Proteases are carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used 10 herein, "protease" means a naturally-occurring protease or a recombinant protease.

Naturally-occurring proteases include  $\alpha$ -aminoacylpeptide hydrolase, peptidylamino acid hydrolase, acylamino hydrolase, serine carboxypeptidase, metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exo-proteases.

Subtilisins are bacterial or fungal proteases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally-occurring subtilisin or a recombinant subtilisin. A series of naturally-occurring subtilisins is known to be produced and often secreted by various microbial species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this 15 series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative 20 order of these amino acids, reading from the amino to carboxy terminus, is aspartate-histidine-serine. In the chymotrypsin related proteases, the relative order, however, is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases. Examples include but are not limited to the 25 subtilisins identified in Fig. 3 herein. Generally and for purposes of the present invention, numbering of the amino acids in proteases corresponds to the numbers assigned to the mature *Bacillus amyloliquefaciens* subtilisin sequence presented in Fig. 1.

"Recombinant subtilisin" or "recombinant protease" refer to a subtilisin or protease in which the DNA sequence encoding the subtilisin or protease is modified to produce a variant (or mutant) DNA sequence which encodes the substitution, deletion or insertion of 30 one or more amino acids in the naturally-occurring amino acid sequence. Suitable

methods to produce such modification, and which may be combined with those disclosed herein, include those disclosed in US Patent 4,760,025 (RE 34,606), US Patent 5,204,015 and US Patent 5,185,258.

"Non-human subtilisins" and the DNA encoding them may be obtained from many 5 procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as *E. coli* or *Pseudomonas* and gram positive bacteria such as *Micrococcus* or *Bacillus*. Examples of eucaryotic organisms from which subtilisin and their genes may be obtained include yeast such as *Saccharomyces cerevisiae*, fungi such as *Aspergillus* sp.

10 "Human subtilisin" means proteins of human origin which have subtilisin type catalytic activity, e.g., the kexin family of human derived proteases. An example of such a protein is represented by the sequence in Fig. 7. Additionally, derivatives or homologs of human subtilisin, including those from non-human sources such as mouse or rabbit, which retain the essential ability to hydrolyze peptide bonds and have at least 50%, preferably at 15 least 65% and most preferably at least 80% homology to the protein of Fig. 7 are considered human subtilisins for the purpose of the invention.

A "protease variant" has an amino acid sequence which is derived from the amino acid sequence of a "precursor protease". The precursor proteases include naturally-occurring proteases and recombinant proteases. The amino acid sequence of the 20 protease variant is "derived" from the precursor protease amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the amino acid sequence of the precursor protease rather than manipulation of the precursor protease enzyme *per se*. Suitable methods for such manipulation of the precursor DNA 25 sequence include methods disclosed herein, as well as methods known to those skilled in the art (see, for example, EP 0 328299, WO89/06279 and the US patents and applications already referenced herein).

The amino acid position numbers used herein refer to those assigned to the mature *Bacillus amyloliquefaciens* subtilisin sequence presented in Fig. 1. The invention, 30 however, is not limited to the mutation of this particular subtilisin but extends to precursor proteases containing amino acid residues at positions which are "equivalent" to the particular identified residues in *Bacillus amyloliquefaciens* subtilisin. In a preferred embodiment of the present invention, the precursor protease is *Bacillus lentus* subtilisin and the substitutions, deletions or insertions are made at the equivalent amino acid 35 residue in *B. lentus* corresponding to those listed above.

A residue (amino acid) of a precursor protease is equivalent to a residue of *Bacillus amyloliquefaciens* subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analogous to a specific residue or portion of that residue in *Bacillus amyloliquefaciens* subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor protease is directly compared to the *Bacillus amyloliquefaciens* subtilisin primary sequence and particularly to a set of residues known to be invariant in subtilisins for which the sequence is known. For example, Fig. 2 herein shows the conserved residues as between *B. amyloliquefaciens* subtilisin and *B. lenthus* subtilisin. After aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of *Bacillus amyloliquefaciens* subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 50% of conserved residues is also adequate to define equivalent residues. Conservation of the catalytic triad, Asp32/His64/Ser221 should be maintained.

For example, the amino acid sequence of subtilisin from *Bacillus amyloliquefaciens*, *Bacillus subtilis*, *Bacillus licheniformis* (*carlsbergensis*) and *Bacillus lenthus* can be aligned to provide the maximum amount of homology between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. The conserved residues as between BPN' and *B. lenthus* are identified in Fig. 2.

These conserved residues, thus, may be used to define the corresponding equivalent amino acid residues of *Bacillus amyloliquefaciens* subtilisin in other subtilisins such as subtilisin from *Bacillus lenthus* (PCT Publication No. W089/06279 published July 13, 1989), the preferred protease precursor enzyme herein, or the subtilisin referred to as PB92 (EP 0 328 299), which is highly homologous to the preferred *Bacillus lenthus* subtilisin. The amino acid sequences of certain of these subtilisins are aligned in Figs. 3A and 3B with the sequence of *Bacillus amyloliquefaciens* subtilisin to produce the maximum homology of conserved residues. As can be seen, there are a number of deletions in the sequence of *Bacillus lenthus* as compared to *Bacillus amyloliquefaciens* subtilisin. Thus, for example, the equivalent amino acid for Val165 in *Bacillus amyloliquefaciens* subtilisin in the other subtilisins is isoleucine for *B. lenthus* and *B. licheniformis*.

Thus, for example, the amino acid at position +170 is lysine (K) in both *B. amyloliquefaciens* and *B. licheniformis* subtilisins and arginine (R) in Savinase. In one embodiment of the protease variants of the invention, however, the amino acid equivalent to +170 in *Bacillus amyloliquefaciens* subtilisin is substituted with aspartic acid (D). The 5 abbreviations and one letter codes for all amino acids in the present invention conform to the PatentIn User Manual (GenBank, Mountain View, CA) 1990, p.101.

"Equivalent residues" may also be defined by determining homology at the level of tertiary structure for a precursor protease whose tertiary structure has been determined by x-ray crystallography. Equivalent residues are defined as those for which the atomic 10 coordinates of two or more of the main chain atoms of a particular amino acid residue of the precursor protease and *Bacillus amyloliquefaciens* subtilisin (N on N, CA on CA, C on C and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the protease in question 15 to the *Bacillus amyloliquefaciens* subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$R \text{ factor} = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

Equivalent residues which are functionally analogous to a specific residue of 20 *Bacillus amyloliquefaciens* subtilisin are defined as those amino acids of the precursor protease which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the *Bacillus amyloliquefaciens* subtilisin. Further, they are those residues of the precursor protease (for which a tertiary structure has been obtained by x-ray crystallography) which occupy an analogous position to the extent that, although the 25 main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie with 0.13nm of the corresponding side chain atoms of *Bacillus amyloliquefaciens* subtilisin. The coordinates of the three dimensional structure of 30 *Bacillus amyloliquefaciens* subtilisin are set forth in EPO Publication No. 0 251 446 (equivalent to US Patent 5,182,204, the disclosure of which is incorporated herein by

reference) and can be used as outlined above to determine equivalent residues on the level of tertiary structure.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a variant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally-occurring sequence. The protease variants of the present invention include the mature forms of protease variants, as well as the pro- and prepro- forms of such protease variants. The prepro- forms are the preferred construction since this facilitates the expression, secretion and maturation of the protease variants.

"Prosequence" refers to a sequence of amino acids bound to the N-terminal portion of the mature form of a protease which when removed results in the appearance of the "mature" form of the protease. Many proteolytic enzymes are found in nature as translational proenzyme products and, in the absence of post-translational processing, are expressed in this fashion. A preferred prosequence for producing protease variants is the putative prosequence of *Bacillus amyloliquefaciens* subtilisin, although other protease prosequences may be used.

A "signal sequence" or "prosequence" refers to any sequence of amino acids bound to the N-terminal portion of a protease or to the N-terminal portion of a proprotease which may participate in the secretion of the mature or pro forms of the protease. This definition of signal sequence is a functional one, meant to include all those amino acid sequences encoded by the N-terminal portion of the protease gene which participate in the effectuation of the secretion of protease under native conditions. The present invention utilizes such sequences to effect the secretion of the protease variants as defined herein. One possible signal sequence comprises the first seven amino acid residues of the signal sequence from *Bacillus subtilis* subtilisin fused to the remainder of the signal sequence of the subtilisin from *Bacillus lentus* (ATCC 21536).

A "prepro" form of a protease variant consists of the mature form of the protease having a prosequence operably linked to the amino terminus of the protease and a "pre" or "signal" sequence operably linked to the amino terminus of the prosequence.

"Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable

mRNA ribosome binding sites and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the 5 genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

The "host cells" used in the present invention generally are prokaryotic or 10 eucaryotic hosts which preferably have been manipulated by the methods disclosed in US Patent 4,760,025 (RE 34,606) to render them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing protease is the *Bacillus* strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in US Patent 15 5,264,366. Other host cells for expressing protease include *Bacillus subtilis* I168 (also described in US Patent 4,760,025 (RE 34,606) and US Patent 5,264,366, the disclosure of which are incorporated herein by reference), as well as any suitable *Bacillus* strain such as *B. licheniformis*, *B. lentus*, etc.

Host cells are transformed or transfected with vectors constructed using 20 recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the protease variants or expressing the desired protease variant. In the case of vectors which encode the pre- or prepro-form of the protease variant, such variants, when expressed, are typically secreted from the host cell into the host cell medium.

25 "Operably linked", when describing the relationship between two DNA regions, simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it 30 controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

The genes encoding the naturally-occurring precursor protease may be obtained in accord with the general methods known to those skilled in the art. The methods generally comprise synthesizing labeled probes having putative sequences encoding regions of the 35 protease of interest, preparing genomic libraries from organisms expressing the protease,

and screening the libraries for the gene of interest by hybridization to the probes.

Positively hybridizing clones are then mapped and sequenced.

The cloned protease is then used to transform a host cell in order to express the protease. The protease gene is then ligated into a high copy number plasmid. This 5 plasmid replicates in hosts in the sense that it contains the well-known elements necessary for plasmid replication: a promoter operably linked to the gene in question (which may be supplied as the gene's own homologous promoter if it is recognized, i.e., transcribed, by the host), a transcription termination and polyadenylation region (necessary for stability of the mRNA transcribed by the host from the protease gene in 10 certain eucaryotic host cells) which is exogenous or is supplied by the endogenous terminator region of the protease gene and, desirably, a selection gene such as an antibiotic resistance gene that enables continuous cultural maintenance of plasmid-infected host cells by growth in antibiotic-containing media. High copy number plasmids also contain an origin of replication for the host, thereby enabling large numbers of 15 plasmids to be generated in the cytoplasm without chromosomal limitations. However, it is within the scope herein to integrate multiple copies of the protease gene into host genome. This is facilitated by procaryotic and eucaryotic organisms which are particularly susceptible to homologous recombination.

In one embodiment, the gene can be a natural gene such as that from *B. lenthus* or 20 *B. amyloliquefaciens*. Alternatively, a synthetic gene encoding a naturally-occurring or mutant precursor protease may be produced. In such an approach, the DNA and/or amino acid sequence of the precursor protease is determined. Multiple, overlapping synthetic single-stranded DNA fragments are thereafter synthesized, which upon hybridization and ligation produce a synthetic DNA encoding the precursor protease. An 25 example of synthetic gene construction is set forth in Example 3 of US Patent 5,204,015, the disclosure of which is incorporated herein by reference.

Once the naturally-occurring or synthetic precursor protease gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond 30 synthesis of the naturally-occurring precursor protease. Such modifications include the production of recombinant proteases as disclosed in US Patent 4,760,025 (RE 34,606) and EPO Publication No. 0 251 446 and the production of protease variants described herein.

The following cassette mutagenesis method may be used to facilitate the 35 construction of the protease variants of the present invention, although other methods may be used. First, the naturally-occurring gene encoding the protease is obtained and

sequenced in whole or in part. Then the sequence is scanned for a point at which it is desired to make a mutation (deletion, insertion or substitution) of one or more amino acids in the encoded enzyme. The sequences flanking this point are evaluated for the presence of restriction sites for replacing a short segment of the gene with an oligonucleotide pool

5 which when expressed will encode various mutants. Such restriction sites are preferably unique sites within the protease gene so as to facilitate the replacement of the gene segment. However, any convenient restriction site which is not overly redundant in the protease gene may be used, provided the gene fragments generated by restriction digestion can be reassembled in proper sequence. If restriction sites are not present at

10 locations within a convenient distance from the selected point (from 10 to 15 nucleotides), such sites are generated by substituting nucleotides in the gene in such a fashion that neither the reading frame nor the amino acids encoded are changed in the final construction. Mutation of the gene in order to change its sequence to conform to the desired sequence is accomplished by M13 primer extension in accord with generally

15 known methods. The task of locating suitable flanking regions and evaluating the needed changes to arrive at two convenient restriction site sequences is made routine by the redundancy of the genetic code, a restriction enzyme map of the gene and the large number of different restriction enzymes. Note that if a convenient flanking restriction site is available, the above method need be used only in connection with the flanking region

20 which does not contain a site.

Once the naturally-occurring DNA or synthetic DNA is cloned, the restriction sites flanking the positions to be mutated are digested with the cognate restriction enzymes and a plurality of end termini-complementary oligonucleotide cassettes are ligated into the gene. The mutagenesis is simplified by this method because all of the oligonucleotides

25 can be synthesized so as to have the same restriction sites, and no synthetic linkers are necessary to create the restriction sites.

In one aspect of the invention, the objective is to secure a variant protease having altered allergenic potential as compared to the precursor protease, since decreasing such potential enables safer use of the enzyme. While the instant invention is useful to lower

30 allergenic potential, the mutations specified herein may be utilized in combination with mutations known in the art to result altered thermal stability and/or altered substrate specificity, modified activity or altered alkaline stability as compared to the precursor.

Accordingly, the present invention is directed to altering the capability of the T-cell

35 epitope which includes residue positions 170-173 in *Bacillus lenthus* to induce T-cell

proliferation. One particularly preferred embodiment of the invention comprises making modification to either one or all of R170D, Y171Q and/or N173D. Similarly, as discussed in detail above, it is believed that the modification of the corresponding residues in any protease will result in a the neutralization of a key T-cell epitope in that protease. Thus, in 5 combination with the presently disclosed mutations in the region corresponding to amino acid residues 170-173, substitutions at positions corresponding to N76D/S103A/V104I/G159D optionally in combination with one or more substitutions selected from the group consisting of positions corresponding to V68A, T213R, A232V, Q236H, Q245R, and T260A of *Bacillus amyloliquefaciens* subtilisin may be used, in 10 addition to decreasing the allergenic potential of the variant protease of the invention, to modulate overall stability and/or proteolytic activity of the enzyme. Similarly, the substitutions provided herein may be combined with mutation at the Asparagine (N) in *Bacillus lenthus* subtilisin at equivalent position +76 to Aspartate (D) in combination with the mutations S103A/V104I/G159D and optionally in combination with one or more 15 substitutions selected from the group consisting of positions corresponding to V68A, T213R, A232V, Q236H, Q245R, and T260A of *Bacillus amyloliquefaciens* subtilisin, to produce enhanced stability and/or enhanced activity of the resulting mutant enzyme.

The most preferred embodiments of the invention include the following specific combinations of substituted residues corresponding to positions:

20 N76D/S103A/V104I/G159D/K170D/Y171Q/S173D;  
V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D /Q236H;  
V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D /Q236H/Q245R;  
V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/A232V/Q236H/Q245R; and  
V68A/N76D//S103A/V104I/G159D/K170D/Y171Q/S173D/T213R/A232V/Q236H/  
25 Q245R/T260A of *Bacillus amyloliquefaciens* subtilisin. These substitutions are preferably made in *Bacillus lenthus* (recombinant or native-type) subtilisin, although the substitutions may be made in any *Bacillus* protease.

Based on the screening results obtained with the variant proteases, the noted mutations noted above in *Bacillus amyloliquefaciens* subtilisin are important to the 30 proteolytic activity, performance and/or stability of these enzymes and the cleaning or wash performance of such variant enzymes.

Many of the protease variants of the invention are useful in formulating various detergent compositions. A number of known compounds are suitable surfactants useful in compositions comprising the protease mutants of the invention. These include nonionic, 35 anionic, cationic, anionic or zwitterionic detergents, as disclosed in US 4,404,128 to Barry

J. Anderson and US 4,261,868 to Jiri Flora, et al. A suitable detergent formulation is that described in Example 7 of US Patent 5,204,015 (previously incorporated by reference). The art is familiar with the different formulations which can be used as cleaning compositions. In addition to typical cleaning compositions, it is readily understood that the protease variants of the present invention may be used for any purpose that native or wild-type proteases are used. Thus, these variants can be used, for example, in bar or liquid soap applications, dishcare formulations, contact lens cleaning solutions or products, peptide hydrolysis, waste treatment, textile applications, as fusion-cleavage enzymes in protein production, etc. The variants of the present invention may comprise, in addition to decreased allergenicity, enhanced performance in a detergent composition (as compared to the precursor). As used herein, enhanced performance in a detergent is defined as increasing cleaning of certain enzyme sensitive stains such as grass or blood, as determined by usual evaluation after a standard wash cycle.

Proteases of the invention can be formulated into known powdered and liquid detergents having pH between 6.5 and 12.0 at levels of about .01 to about 5% (preferably .1% to .5%) by weight. These detergent cleaning compositions can also include other enzymes such as known proteases, amylases, cellulases, lipases or endoglycosidases, as well as builders and stabilizers.

The addition of proteases of the invention to conventional cleaning compositions does not create any special use limitation. In other words, any temperature and pH suitable for the detergent is also suitable for the present compositions as long as the pH is within the above range, and the temperature is below the described protease's denaturing temperature. In addition, proteases of the invention can be used in a cleaning composition without detergents, again either alone or in combination with builders and stabilizers.

The variant proteases of the present invention can be included in animal feed such as part of animal feed additives as described in, for example, US 5,612,055; US 5,314,692; and US 5,147,642.

One aspect of the invention is a composition for the treatment of a textile that includes variant proteases of the present invention. The composition can be used to treat for example silk or wool as described in publications such as RD 216,034; EP 134,267; US 4,533,359; and EP 344,259.

The following is presented by way of example and is not to be construed as a limitation to the scope of the claims.

The variants can be screened for proteolytic activity according to methods well known in the art. Preferred protease variants include multiple substitutions at positions corresponding to: N76D/S103A/V104I/G159D/K170D/Y171Q/S173D; V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/Q236H; 5 V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/Q236H/Q245R; V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/A232V/Q236H/Q245R; and V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/T213R/A232V/Q236H/Q245R/T2 60A of *Bacillus amyloliquefaciens* subtilisin.

All publications and patents referenced herein are hereby incorporated by 10 reference in their entirety.

## **EXAMPLES**

### **Example 1**

#### **Assay for the Identification of Peptide T-Cell Epitopes Using Naïve Human T-Cells**

Fresh human peripheral blood cells were collected from "naïve" humans, i.e., persons not known to be exposed to or sensitized to *Bacillus lentinus* protease, for determination of antigenic epitopes in protease from *Bacillus lentinus* and human subtilisin. Naïve humans is intended to mean that the individual is not known to have been exposed 20 to or developed a reaction to protease in the past. Peripheral mononuclear blood cells (stored at room temperature, no older than 24 hours) were prepared for use as follows: Approximately 30 mls of a solution of buffy coat preparation from one unit of whole blood was brought to 50 ml with Dulbecco's phosphate buffered solution (DPBS) and split into two tubes. The samples were underlaid with 12.5 ml of room temperature lymphoprep 25 density separation media (Nycomed density 1.077 g/ml). The tubes were centrifuged for thirty minutes at 600G. The interface of the two phases was collected, pooled and washed in DPBS. The cell density of the resultant solution was measured by hemocytometer. Viability was measured by trypan blue exclusion.

From the resulting solution, a differentiated dendritic cell culture was prepared from 30 the peripheral blood mononuclear cell sample having a density of  $10^8$  cells per 75 ml culture flask in a solution as follows:

- (1) 50 ml of serum free AIM V media (Gibco) was supplemented with a 1:100 dilution beta-mercaptoethanol (Gibco). The flasks were laid flat for two hours at 37°C in 5% CO<sub>2</sub> to allow adherence of monocytes to the flask wall.

(2) Differentiation of the monocyte cells to dendritic cells was as follows: nonadherent cells were removed and the resultant adherent cells (monocytes) combined with 30 ml of AIM V, 800 units/ml of GM-CSF (Endogen) and 500 units/ml of IL-4 (Endogen); the resulting mixture was cultured for 5 days  
5 under conditions at 37°C in 5% CO<sub>2</sub>. After five days, the cytokine TNFα (Endogen) was added to 0.2 units/ml, and the cytokine IL-1α (Endogen) was added to a final concentration of 50 units/ml and the mixture incubated at 37°C in 5% CO<sub>2</sub> for two more days.

10 (3) On the seventh day, Mitomycin C was added to a concentration of 50 microgram/ml was added to stop growth of the now differentiated dendritic cell culture. The solution was incubated for 60 minutes at 37°C in 5% CO<sub>2</sub>. Dendritic cells were collected by gently scraping the adherent cells off the bottom of the flask with a cell scraper. Adherent and non-adherent cells were then centrifuged at 600G for 5 minutes, washed in DPBS and counted.

15 (4) The prepared dendritic cells were placed into a 96 well round bottom array at 2x10<sup>4</sup>/well in 100 microliter total volume of AIM V media.

CD4+ T cells were prepared from frozen aliquots of the peripheral blood cell samples used to prepare the dendritic cells using the human CD4+ Cellect Kit (Biotex) as per the manufacturers instructions with the following modifications: the aliquots were  
20 thawed and washed such that approximately 10<sup>8</sup> cells will be applied per Cellect column; the cells were resuspended in 4 ml DPBS and 1 ml of the Cell reagent from the Cellect Kit, the solution maintained at room temperature for 20 minutes. The resultant solution was centrifuged for five minutes at 600G at room temperature and the pellet resuspended in 2 ml of DPBS and applied to the Cellect columns. The effluent from the columns was  
25 collected in 2% human serum in DPBS. The resultant CD4+ cell solution was centrifuged, resuspended in AIMV media and the density counted.

The CD4+ T-cell suspension was resuspended to a count of 2x10<sup>6</sup>/ml in AIM V media to facilitate efficient manipulation of the 96 well plate.

Peptide antigen is prepared from a 1M stock solution in DMSO by dilution in AIM V  
30 media at a 1:10 ratio. 10 microliters of the stock solution is placed in each well of the 96 well plate containing the differentiated dendritic cells. 100 microliter of the diluted CD4+ T-cell solution as prepared above is further added to each well. Useful controls include diluted DMSO blanks, and tetanus toxoid positive controls.

The final concentrations in each well, at 210 microliter total volume are as follows:

- 2x10<sup>4</sup> CD4+
- 2x10<sup>5</sup> dendritic cells (R:S of 10:1)
- 5 mM peptide

5

### Example 2

#### Identification of T-Cell Epitopes in Protease from *Bacillus lenthus* and Human subtilisin

Peptides for use in the assay described in Example 1 were prepared based on the  
10 *Bacillus lenthus* and human subtilisin amino acid sequence. Peptide antigens were  
designed as follows. From the full length amino acid sequence of either human subtilisin  
or *Bacillus lenthus* protease provided in Figure 1, 15mers were synthetically prepared, each  
15mer overlapping with the previous and the subsequent 15mer except for three residues.

Peptides used correspond to amino acid residue strings in *Bacillus lenthus* as  
15 provided in Figure 8, and peptides correspond to amino acid residues in human subtilisin  
as provided in Figure 7. The peptides used corresponding to the proteases is provided in  
Fig. 6. All tests were performed at least in duplicate. All tests reported displayed robust  
positive control responses to the antigen tetanus toxoid. Responses were averaged within  
each experiment, then normalized to the baseline response. A positive event was  
20 recorded if the response was at least 3 times the baseline response.

The immunogenic response (i.e., T-cell proliferation) to the prepared peptides from  
human subtilisin and *Bacillus lenthus* was tallied and is provided in Figures 4 and 5,  
respectively. T-cell proliferation was measured by the incorporated tritium method. The  
results shown in Figures 4 and 5 as a comparison of the immunogenic additive response  
25 in 10 individuals (Figure 4) and 16 individuals (Figure 5) to the various peptides.  
Response is indicated as the added response wherein 1.0 equals a baseline response for  
each sample. Thus, in Figure 4, a reading of 10.0 or less is the baseline response and in  
Figure 5 a reading of 16.0 or less the baseline response.

As indicated in Figures 4 and 5, the immunogenic response of the naïve blood  
30 samples from unsensitized individuals showed a marked allergic response at the  
peptide fragment from *Bacillus lenthus* corresponding to residues 170-173 of *Bacillus*  
*amyloliquefaciens* protease. As expected, the corresponding fragment in human subtilisin  
evokes merely baseline response.

Fig. 9 shows the T-cell response to peptides derived from *Bacillus lenthus* protease  
35 in a sample taken from an individual known to be hypersensitive to *Bacillus lenthus*

protease. Peptide E05 represents the region corresponding to 170-173 in protease from *Bacillus amyloliquefaciens*. As shown in Fig. 9, the hypersensitive individual was highly responsive to the T-cell epitope represented by the peptide E05. This result confirms that, by practicing the assay according to the invention, it is possible to predict the major epitopes identified by the T-cells of a hypersensitive individual.

Fig. 10 shows the T-cell response to various alanine substitutions in the E05 peptide derived from *Bacillus lenthus* protease in a sample taken from an individual known to be hypersensitive to *Bacillus lenthus* protease. Alanine substitutions were used as substitutions for the purpose of determining the role of any specific residue within the epitope. The legend of Figure 10 refers to the position of the peptide in which an alanine was substituted, i.e., in peptide E06 (sequence GSISYPARYANAMAV), G to A = 2, S to A = 3, I to A = 4, S to A = 5, Y to A = 6, P to A = 7, R to A = 8, Y to A = 9, N to A = 10, M to A = 11 and V to A = 12. As indicated in Figure 10, substitution of either of the residues R170A, Y171A and/or N173A in protease from *Bacillus lenthus* results in dramatically reduced response in the hypersensitive individual's blood sample.

From these results, it is apparent that the residues 170, 171 and 173 are critical for T-cell response within this peptide. Accordingly, it is further apparent that these residues are largely responsible for the initiation of allergic reaction within the protease from *Bacillus lenthus*.

**WE CLAIM:**

1. A protease variant comprising a substitution made at one or more of positions in a precursor protease corresponding to K170D, Y171Q and/or S173D of *Bacillus amyloliquefaciens* subtilisin.

5        2. The protease variant according to claim 1, further comprising a substitution at one or more positions in a precursor protease equivalent to those selected from the group consisting of N76D, S103A, V104I, G159D, V68A, T213R, A232V, Q236H, Q245R, and T260A.

10      3. The protease variant according to claim 2 which is derived from a *Bacillus subtilisin*.

4. The protease variant according to claim 3 which is derived from *Bacillus lentus* subtilisin or *Bacillus amyloliquefaciens* subtilisin.

5. A DNA encoding a protease variant of claim 1.

6. An expression vector encoding the DNA of claim 5.

15      7. A host cell transformed with the expression vector of claim 6.

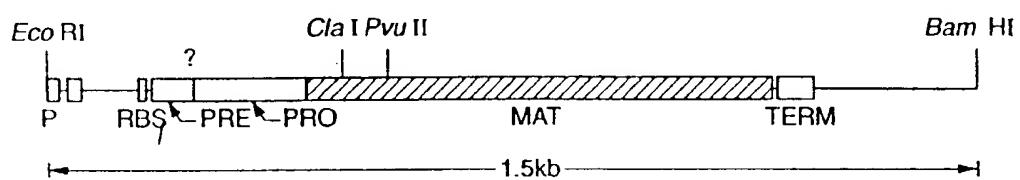
8. A cleaning composition comprising the protease variant of claim 1.

9. An animal feed comprising the protease variant of claim 1.

10. A composition for treating a textile comprising the protease variant of claim 1.

15      11. A protease variant according to claim 1, comprising combined substitution sets  
20     selected from the group consisting of positions corresponding to K170D/Y171Q/S173D;  
N76D/S103A/V104I/G159D/ K170D/Y171Q/S173D; V68A/N76D/S103A/V104I/G159D/  
K170D/Y171Q/S173D/Q236H; V68A/N76D/S103A/V104I/G159D/  
K170D/Y171Q/S173D/Q236H/Q245R;  
V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/  
25     A232V/Q236H/Q245R; and  
V68A/N76D/S103A/V104I/G159D/K170D/Y171Q/S173D/T213R/A232V/Q236H/  
Q245R/T260A of *Bacillus amyloliquefaciens* subtilisin.

12. A method for determining T-cell epitopes in humans comprising the steps of:
  - (a) obtaining from a single blood source a solution of dendritic cells and a solution of naïve CD4+ and/or CD8+ T-cells;
  - (b) promoting differentiation in said solution of dendritic cells;
  - 5 (c) combining said solution of differentiated dendritic cells and said naïve CD4+ and/or CD8+ T-cells with a peptide of interest;
  - (d) measuring the production of antibodies in said step (c).
13. A method of reducing the allergenicity of a protein comprising the steps of:
  - (a) identifying a T-cell epitope in said protein;
  - 10 (b) modifying said protein to neutralize said T-cell epitope.
14. The method according to claim 13, wherein said epitope is modified by:
  - (a) substituting the amino acid sequence of the epitope with an analogous sequence from a human homolog to the protein of interest;
  - (b) substituting the amino acid sequence of the epitope with an analogous sequence from a non-human homolog to the protein of interest, which analogous sequence produces a lesser allergenic response from T-cells than that of the protein of interest; or
  - 15 (c) substituting the amino acid sequence of the epitope with a sequence which substantially mimics the major tertiary structure attributes of the epitope, but which produces a lesser allergenic response from T-cells than that of the protein of interest.
- 20 15. A protein having reduced allergenicity made by the method according to claim 14.
16. A protein having reduced allergenicity, wherein said protein comprises a modification comprising the substitution or deletion of amino acid residues which are identified as within a T-cell epitope according to the assay provided in claim 13.



**FIG.\_1A**

2/16

5' UTR and RBS sequence:

```

    1   2   3   4   5
    |   |   |   |   |
  AAG GGC AAG AAA GTC TGA ATC AGT TGT CTC TTT GCT TTA ATC TTT ACG ATG GCG TCC GGC AGC ACA TCC
  |   |   |   |   |
  AAG Lys Val Thr Ile Ser Leu Ala Phe Ala Leu Ala Leu Phe Thr Met Ala Phe Gly Ser Thr Ser
  |   |   |   |   |
  AGA GGC AAA TAA TGG ATC AAC TCA GAA AGG GAA AAA TAT ATT GTC GGG TTT AAA CAG ACA ATG AGC ACG ATG
  |   |   |   |   |
  AG Glys Val Lys Ser Asn Glu Glu Lys Tyr Ile Val Gly Lys Glu Phe Lys Tyr Val Asp Ala
  |   |   |   |   |
  AGC GGC GCT AAG AAG AAA GAT GTC ATT TCT GAA AAA GGC GGG AAA TAT ATT GTC GGG AAA GTG CAA AAG CAA TAC
  |   |   |   |   |
  Ser Ala Ala Lys Lys Asp Val Ile Ser Glu Lys Gly Lys Val Lys Asp Pro Ser Val Ala Tyr Val Glu Asp
  |   |   |   |   |
  Ser Ala Ala Lys Lys Asp Val Ile Ser Glu Lys Gly Lys Val Lys Asp Pro Ser Val Ala Tyr Val Glu Asp
  |   |   |   |   |
  GCT TCA GCT ACA TTA AAC GAA AAA GAT GTC ATT TCT GAA AAA GGC GGG AAA TAT ATT GTC GGG AAA GTG CAA AAG CAA TAC
  |   |   |   |   |
  -30   -20   -10   10   20
  Ala Ser Ala Thr Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Asp Pro Ser Val Ala Tyr Val Glu Asp
  |   |   |   |   |
  -324   -315   -305   -295   -285
  Ala Ser Ala Thr Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Asp Pro Ser Val Ala Tyr Val Glu Asp
  |   |   |   |   |
  CAC GCA CAT GCG TAC GCG CAG TCC GTG CCT TAC GCC GTA TCA CAA ATT AAA GCC CCT GCT CTG CAC TCT CAA
  |   |   |   |   |
  His Val Ala His Ala Tyr Ala Gln Ser Val Pro Tyr Gly Val Ser Glu Ile Lys Ala Pro Ala Leu His Ser Glu
  |   |   |   |   |
  399   389   379   369   359
  Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys Val
  |   |   |   |   |
  474   464   454   444   434
  GGC TAC ACT GGA TCA AAT GTC TTA AAA GTC GCG GTC GAT GAC ACC GGT ATC GAT TCT CAT CCT GAT TTA AAG GTC
  |   |   |   |   |
  40   30   20   10   0
  
```

FIG.-1B - 1

3/16

549	Ala	Gly	Gly	Ala	Ser	Mel	Val	Pro	Ser	Glu	Thr	Asn	Pro	Phe	Gln	Asp	Asn	Ser	His	Gly	Thr	His	Val	Ala	
600	GCC	GGC	GGA	GCC	GGC	AGC	ATG	GTT	CCT	TCT	GAA	ACA	AAT	CCT	TTC	CAA	GAC	AAC	TCT	CAC	GGA	ACT	CAC	GTT	GCC
669	Val	Leu	Gly	Ala	Asp	Gly	Ser	Gly	Gln	Tyr	Ser	Ile	Ile	Gly	Ile	Glu	Trp	Ala	Leu	Tyr	Ala	Val	Lys	Ala	
728	Asp	Val	Ile	Asn	Mel	Ser	Leu	Gly	Gly	Pro	Ser	Gly	Ser	Ala	Leu	Lys	Ala	Val	Asp	Lys	Ala	Val	Ala	GCC	
787	Ser	Gly	Val	Val	Val	Ala	Ala	Gly	Asn	Glu	Gly	Asn	Glu	Gly	Ser	Ser	Ser	Thr	Val	Gly	Tyr	Pro	Gly	CCT	
849	Lys	Tyr	Pro	Ser	Val	Ile	Ala	Val	Gly	Ala	Val	Asp	Ser	Ser	Gln	Arg	Ala	Ser	Phe	Ser	Val	Gly	Pro	Pro	
908	AAA	TAC	TCT	TCT	GTC	ATT	GCA	GTA	GGC	GCT	GTT	GAC	AGC	AAC	CAA	AGA	GCA	TCT	TTC	TCA	AGC	GGA	CCT	ATC	
967	Glu	Leu	Asp	Val	Met	Ala	Pro	Gly	Val	Ser	Ile	Gln	Ser	Thr	Leu	Pro	Gly	Asn	Lys	Tyr	Gly	Ala	Tyr	Asn	
1027	GAG	CTT	GAT	GTC	ATG	GCA	CCT	GGC	GTA	TCT	ATC	CAA	AGC	ACG	CTT	CCT	GGG	GGG	TAC	AAC	AAA	TAC	GCC	TAC	

**FIG. - 1B - 2**

1149 CAA GTC CCC AGC AGT TTA GAA AAC ACC ACT ACA AAA CTT GGT GAT TCT GGC TAC TAT GGA AAA GGG CTG ATC AAC  
 270 Val Arg Ser Ser Leu Glu Asn Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Lys Gly Leu Ile Asn

275 Val Gln Ala Ala Gln OC TERM  
 1224 GTA CAG GCG GCA GCT CAG TAA AACATAAAAACCGGCC TTGGCCCCGGGTTTTATTTCTCCGCAAGTCAATTGGCTC

1316 ATAATCGACGGATGGCTCCCTCTGAAATTAAACGAGAAACGGGAACTTGACCCGGCTAGTCCGTAAAGTCCGAAACTCGTCAATCGGATC

### *FIG.-1B - 3*

*FIG.-1B - 1*

*FIG.-1B - 2*

*FIG.-1B - 3*

### *FIG.-1B*

CONSERVED RESIDUES IN SUBTILISINS FROM  
*BACILLUS AMYLOLIQUEFACIENS*

1	10	20
A Q S V P . G . . . .	A P A . H . . G	
21	30	40
. T G S . V K V A V . D . G . . . .	H P	
41	50	60
D L . . . G G A S . V P . . . . .	Q D	
61	70	80
. N . H G T H V A G T . A A L N N S I G		
81	90	100
V L G V A P S A . L Y A V K V L G A . G		
101	110	120
S G . . S . L . . G . E W A . N . . .		
121	130	140
V . N . S L G . P S . S . . . . A . .		
141	150	160
. . . . . G V . V V A A . G N . G . . .		
161	170	180
. . . . . Y P . . Y . . . . A V G A .		
181	190	200
D . . N . . A S F S . . G . . L D . . A		
201	210	220
P G V . . Q S T . P G . . Y . . . N G T		
221	230	240
S M A . P H V A G A A L . . . K . . .		
241	250	260
W . . . Q . R . . L . N T . . . L G . . .		
261	270	
. . Y G . G L . N . . A A . .		

**FIG.\_2**

6/16

## COMPARISON OF SUBUTILISIN SEQUENCES FROM:

*B. amyloliquefaciens**B. subtilis**B. licheniformis**B. lentus*

01	10	20	30	
A Q S V P Y G V S Q I K A P A L H S Q G Y T G S N V R V A V I D S S G I D S S H P	A Q S V P Y G I S Q I K A P A L H S Q G Y T G S N V R V A V I D S S G I D S S H P	A Q T V P Y G I S P L I K A D K V Q A Q G F K G A N V R G L T G S G V R V A V L D T G I Q A S H P	A Q S V P W G I S R V Q A P A H N R G L T G S G V R V A V L D T G I S T * H P	
41				
41	50	60	70	
D L K V A G G A S M V P S E T N P P Q D N N S H G T H V A G T V A A L N N S I G	D L N V R G G A S P V P S E T N P Y Q D G S S H G T H V A G T I A A L N N S I G	D L N V V G G A S P V A G E A Y N * T D G N G H G T H V A G T V A A L D N T T G	D L N I R G G A S P V P G E * P S T Q D G N G H G T H V A G T I A A L N N S I G	
81				
81	90	100	110	
V L G V A P S A S S L Y A V K V L G A D G S G Q Y S W I I N G I E W A I A N N M D	V L G V S P S A S S L Y A V K V L D S T C G S G Q Y S W I I N G I E W A I S N N M D	V L G V A P S V S S L Y A V K V L N S S G S G S Y S S G I V S G I E W A T T N G M D	V L G V A P S A E L Y A V K V L G A S G S G S V S S I A Q G L E W A G N N G M H	
121				
121	130	140	150	
V I N M S L G G P S G S S A A L K A A V D K A V A S G V V V A A A G N E G T S G	V I N M S L G G P P T G S S T A L K T V V D K A V S S G I V V V A A A A G N E G S S G	V I N M S L G G A S G S T A M K Q A V D N A Y A R G V V V A A A A G N S G N S G	V A N L S L G S P S P S A T L E Q A V N S A T S R G V L V V A A S G N S G A G S	

**FIG.-3A**

7/16

<p>161                    170                    180                    190</p> <p>S S S T V G Y P G K Y P S V I A V G A V D S S N Q R A S F S S V G P E L D V M A      S T S T V G Y P A K Y P S T I A V G A V N S S N Q R A S F S S A G S E L D V M A      S T N T I G Y P A K Y D S V I A V G A V D S N S N R A S F S S V G A E L E V M A      * * * * I S Y P A R Y N A M A V G A T D Q N N R A S F S Q Y G A G L D I V A</p>	<p>201                    210                    220                    230</p> <p>P G V S I Q S T L P C G N K Y G A Y N G T S M A S P H V A G A A A L I L S K H P N      P G V S I Q S T L P G G G T Y G A Y N G T S M A T P H V A G A A A L I L S K H P T      P G A G V Y S T Y P T N T Y A T L N G T S M A S P H V A G A A A L I L S K H P N      P G V N V Q S T Y P G S T Y A S L N G T S M A T P H V A G A A A L V K Q K N P S</p>	<p>241                    250                    260                    270</p> <p>W T N T Q V R S S L E N T T K L G D S P Y Y G K G L I N V Q A A A Q      W T N A Q V R D R L E S T A T Y L G N S F Y Y G K G L I N V Q A A A Q      L S A S Q V R N R L S S T A T Y L G S S F Y Y G K G L I N V E A A A Q      W S N V Q I R N H L K N T A T S L G S T N L Y G S G L V N A E A A T R</p>
---	---	---

F/G.-3B

311

FIG. 3B

8/16

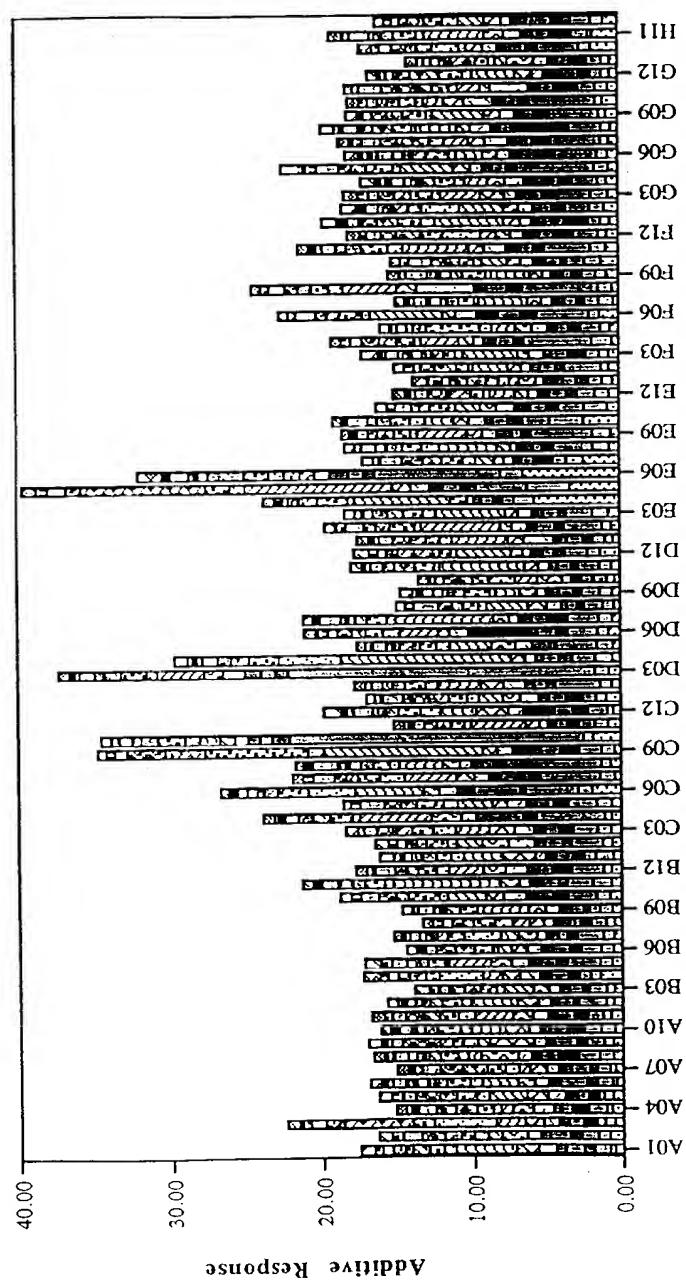


FIG. 4

9/16

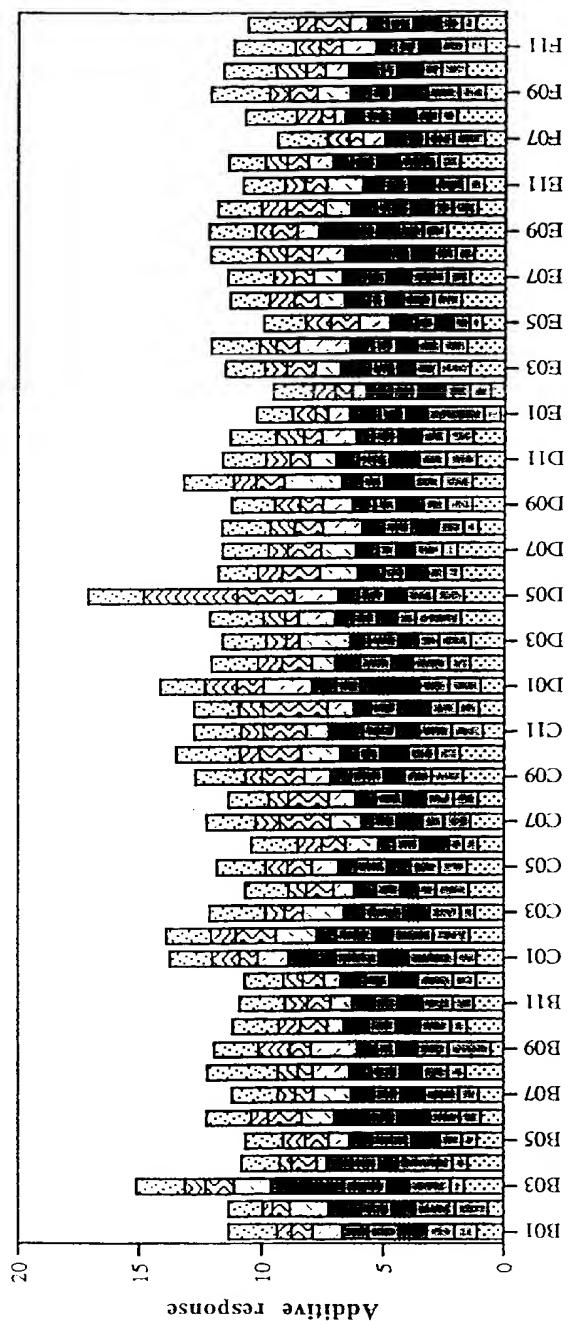


FIG. 5

1	A12	I K D F H V Y F R E S R D A G	49	E12	S A T S R G V L V V A A S G N
2	A11	L E Q A V N S A T S R G V L V	50	E11	S R G V L V V A A S G N S G A
3	A10	A Q S V P W G I S R V Q A P A	51	E10	V L V V A A S G N S G A G S I
4	A9	V P W G I S R V Q A P A A H N	52	E9	V A A S G N S G A G S I S Y P
5	A8	G I S R V Q A P A A H N R G L	53	E8	S G N S G A G S I S Y P A R Y
6	A7	R V Q A P A A H N R G L T G S	54	E7	S G A G S I S Y P A R Y A N A
7	A6	A P A A H N R G L T G S G V K	55	E6	G S I S Y P A R Y A N A M A V
8	A5	A H N R G L T G S G V K V A V	56	E5	S Y P A R Y A N A M A V G A T
9	A4	R G L T G S G V K V A V L D T	57	E4	A R Y A N A M A V G A T D Q N
10	A3	T G S G V K V A V L D T G I S	58	E3	A N A M A V G A T D Q N N N R
11	A2	G V K V A V L D T G I S T H P	59	E2	M A V G A T D Q N N N R A S F
12	A1	V A V L D T G I S T H P D L N	60	E1	G A T D Q N N N R A S F S Q Y
13	B12	L D T G I S T H P D L N I R G	61	F12	D Q N N N R A S F S Q Y G A G
14	B11	G I S T H P D L N I R G G A S	62	F11	N N R A S F S Q Y G A G L D I
15	B10	T H P D L N I R G G A S F V P	63	F10	A S F S Q Y G A G L D I V A P
16	B9	D L N I R G G A S F V P G E P	64	F9	S Q Y G A G L D I V A P G V N
17	B8	I R G G A S F V P G E P S T Q	65	F8	G A G L D I V A P G V N V Q S
18	B7	G A S F V P G E P S T Q D G N	66	F7	L D I V A P G V N V Q S T Y P
19	B6	F V P G E P S T Q D G N G H G	67	F6	V A P G V N V Q S T Y P G S T
20	B5	G E P S T Q D G N G H G T H V	68	F5	G V N V Q S T Y P G S T Y A S
21	B4	S T Q D G N G H G T H V A G T	69	F4	V Q S T Y P G S T Y A S L N G
22	B3	D G N G H G T H V A G T I A A	70	F3	T Y P G S T Y A S L N G T S M
23	B2	G H G T H V A G T I A A L N N	71	F2	G S T Y A S L N G T S M A T P
24	B1	T H V A G T I A A L N N S I G	72	F1	Y A S L N G T S M A T P H V A
25	C12	A G T I A A L N N S I G V L G	73	G12	L N G T S M A T P H V A G A A
26	C11	I A A L N N S I G V L G V A P	74	G11	T S M A T P H V A G A A A L V
27	C10	L N N S I G V L G V A P S A E	75	G10	A T P H V A G A A A L V K Q K
28	C9	S I G V L G V A P S A E L Y A	76	G9	H V A G A A A L V K Q K N P S
29	C8	V L G V A P S A E L Y A V K V	77	G8	G A A A L V K Q K N P S W S N
30	C7	V A P S A E L Y A V K V L G A	78	G7	A L V K Q K N P S W S N V Q I
31	C6	S A E L Y A V K V L G A S G S	79	G6	K Q K N P S W S N V Q I R N H
32	C5	L Y A V K V L G A S G S G S V	80	G5	N P S W S N V Q I R N H L K N
33	C4	V K V L G A S G S G S V S S I	81	G4	W S N V Q I R N H L K N T A T
34	C3	L G A S G S G S V S S I A Q G	82	G3	V Q I R N H L K N T A T S L G
35	C2	S G S G S V S S I A Q G L E W	83	G2	R N H L K N T A T S L G S T N
36	C1	G S V S S I A Q G L E W A G N	84	G1	L K N T A T S L G S T N L Y G
37	D12	S S I A Q G L E W A G N N G M	85	H12	T A T S L G S T N L Y G S G L
38	D11	A Q G L E W A G N N G M H V A	86	H11	S L G S T N L Y G S G L V N A
39	D10	L E W A G N N G M H V A N L S	87	H10	S T N L Y G S G L V N A E A A
40	D9	A G N N G M H V A N L S L G S	88	H9	N L Y G S G L V N A E A A T R
41	D8	N G M H V A N L S L G S P S P			
42	D7	H V A N L S L G S P S P S A T			
43	D6	N L S L G S P S P S A T L E Q			
44	D5	L G S P S P S A T L E Q A V N			
45	D4	P S P S A T L E Q A V N S A T			
46	D3	S A T L E Q A V N S A T S R G			
47	D2	L E Q A V N S A T S R G V L V			
48	D1	A V N S A T S R G V L V V A A			

FIG. 6A

1	A12	IKDFHVYFRESRDA	49	E12	KKIDVLNLSIGGPDF
2	A11	DAELHIFRVFTNNQV	50	E11	DVLNLSIGGPDFMDH
3	A10	PLRRASLSLGSFWH	51	E10	NLSIGGPDFMDHPFV
4	A9	RASLSLGSGFWHATG	52	E9	IGGPDFMDHPFVDKV
5	A8	LSLGSGFWHATGRHS	53	E8	PDFMDHPFVDKVWEL
6	A7	GSGFWHATGRHSSRR	54	E7	MDHPFVDKVWELTAN
7	A6	FWHATGRHSSRRLLR	55	E6	PFVDKVWELTANNVI
8	A5	ATGRHSSRRLLRAIP	56	E5	DKWELTANNVIMVS
9	A4	RHSSRRLLRAIPRQV	57	E4	WELTANNVIMVSAIG
10	A3	SRRLLRAIPRQVAQT	58	E3	TANNVIMVSAIGNDG
11	A2	LLRAIPRQVAQTLQA	59	E2	NVIMVSAIGNDGPLY
12	A1	AIPRQVAQTLQADVL	60	E1	MVSAIGNDGPLYGTJ
13	B12	RQVAQTLQADVLWQM	61	F12	AIGNDGPLYGTLNNP
14	B11	AQTLQADVLWQMGT	62	F11	NDGPLYGTLNNPADQ
15	B10	LQADVLWQMGTGAN	63	F10	PLYGTLNNPADQMDV
16	B9	DVLWQMGTGANVRV	64	F9	GTLNNPADQMDVIGV
17	B8	WQMGTGANVRVAVF	65	F8	NNPADQMDVIGVGGI
18	B7	GYGANVRVAVFDTG	66	F7	ADQMDVIGVGGIDFE
19	B6	GANVRVAVFDTGLSE	67	F6	MDVIGVGGIDFEDNI
20	B5	VRVAVFDTGLSEKHP	68	F5	IGVGGIDFEDNIARF
21	B4	AVFDTGLSEKHPFK	69	F4	GGIDFEDNIARFSSR
22	B3	DTGLSEKHPFKNVK	70	F3	DFEDNIARFSSRGMT
23	B2	LSEKHPFKNVKERT	71	F2	DNIARFSSRGMTTWE
24	B1	KHPFKNVKERTNWT	72	F1	ARFSSRGMTTWEPLG
25	C12	HFKNVKERTNWTNER	73	G12	SSRGMTTWEPLGGYG
26	C11	NVKERTNWTNERLTD	74	G11	GMTTWELPGGYGRMK
27	C10	ERTNWTNERLDDGL	75	G10	TWELPGGYGRMKPDI
28	C9	NWTNERLDDGLGHG	76	G9	LPGGYGRMKPDIVTY
29	C8	NERLDDGLGHGTFV	77	G8	GYGRMKPDIVTYGAG
30	C7	TLDDGLGHGTFVAGV	78	G7	RMKPDIVTYGAGVRG
31	C6	DGLGHGTFVAGVIAS	79	G6	PDIVTYGAGVRGSGV
32	C5	GHGTFVAGVIASMRE	80	G5	VTYGAGVRGSGVKGG
33	C4	TFVAGVIASMRECQG	81	G4	GAGVRGSGVKGGCRA
34	C3	AGVIASMRECQGFAP	82	G3	VRGSGVKGGCRALSG
35	C2	IASMRECQGFAPDAE	83	G2	SGVKGGCRALSGTSV
36	C1	MRECQGFAPDAELHI	84	G1	KGGCRALSGTSVASP
37	D12	CQGFAPDAELHI FRV	85	H12	CRALSGTSVASPVA
38	D11	FAPDAELHI FRVFTN	86	H11	LSGTSVASPVVAGAV
39	D10	DAELHIFRVFTNNQV	87	H10	TSVASPVVAGAVTLL
40	D9	LHIFRVFTNNQVSYT	88	H9	ASPVVAGAVTLLVST
41	D8	FRVFTNNQVSYTWSF	89	H8	VVAGAVTLLVSTVQK
42	D7	FTNNQVSYTWSFLDA	90	H7	GAVTLLVSTVQKREL
43	D6	NQVSYTWSFLDAFN	91	H6	TLLVSTVQKRELVNP
44	D5	SYTWSFLDAFNAIL	92	H5	VSTVQKRELVNPASM
45	D4	SWFLDAFNAILKKI	93	H4	VQKRELVNPASMKQA
46	D3	LDAFNAILKKIDVL	94	H3	RELVNPASMKQALIA
47	D2	FNYAILKKIDVLNLS	95	H2	VNPASMKQALIASAR
48	D1	AILKKIDVLNLSIGG	96	H1	ASMKQALIASARRLP

FIG. 6B

12/16

97	I12	IKDFHVYFRESRDAG
98	I11	DAELHIFRVFTNNQV
99	I10	KQALIASARRLPGVN
100	I9	LIASARRLPGVNMFE
101	I8	SARRLPGVNMFEQGH
102	I7	RLPGVNMFEQGHGKL
103	I6	GVNMFEQGHGKLDLL
104	I5	MFEQGHGKLDLLRAY
105	I4	QGHGKLDLLRAYQIL
106	I3	GKLDLLRAYQILNSY
107	I2	DLLRAYQILNSYKPQ
108	I1	RAYQILNSYKPQASL
109	J12	QILNSYKPQASLSPS
110	J11	NSYKPQASLSPSYID
111	J10	KPQASLSPSYIDLTE
112	J9	ASLSPSYIDLTECPY
113	J8	SPSYIDLTECPYMWP
114	J7	YIDLTECPYMWPYCS
115	J6	LTECPYMWPYCSQPI
116	J5	CPYMWPYCSQPIYYG

FIG. 6C

13/16

MKLVNIWLLLUVLLCGKKHLGDRLEKKSFEKAPCPGCSHTLKVEFSSTVVEYEYIVAFNGYFT  
AKARNSFISSALKSSEVDNWRIIPRNNPSSDYPDFEVIQIKEKQAGLLTLEDHPNIKRVTPQR  
KVFRSLKYAESDPTVPCNETRWSQWKQSSRPLRRASLSLGSGFWHATGRHSSRLLRAIPRQVAQ  
TLQADVLWQMGTGANVRVAVFDTGLSEKHPHFKNVKERTNWTNERTLDDGLGHGTFVAGVIASM  
RECQGFAPDAELHIFRVFTNNQVSYTWFDAFNAYAILKIDVLNLSIGGPDFMDHPFVDKWEL  
TANNVIMVSAIGNDGPLYGTLLNPADQMDVIGVGGIDFEDNIARFSSRGMTTWELPGGYGRMKPD  
IVTYGAGVRGSGVKGGCRALSGTSVASPVAGAVTLLVSTVQKRELVNPASMKQALIASARRLPG  
VNMFEQGHGKLDLLRAYQILNSYKPQASLSPSYIDLTECPYMWYCSPQIYYGGMPTVVNTILN  
GMGVTGRIVDKPDWQPYLPQNGDNIEVAFSYSSVLWPWSGYLAISISVTKKAASWEGLAQGHVMI  
TVASPAETESKNGAEQTSTVKLPIKVLIIPTPRSKRLWDQYHNLRYPFGYFPRDNLRMKNDPL  
DWNGDHIHTNFRDMYQHLLRSMGYFVEVLGAPFTCFDASQYGTLLMVDSEEEYFPEEIAKLLRDVD  
NGLSLVIFSDWYNTSVMRKVKFYDENTRQWWMPDTGGANIPALNELLSVWNMGFSDGLYEGEFTL  
ANHDMDYYASGCSIAKFPEDGVVITQTFKDQGLEVLKQETAVVENVPILGLYQIPAEGGGRIVLYG  
DSNCLDDSHRQKDCFWLLDALLQYTSYGVTPPSLSHSGNRQRPPSGAGSVTPERMENHLLHRYSK  
VLEAHLGDPKPRPLPACPRLSWAKPQPLNETAPSNLWKHQKLLSIDLDKVVLPNFRSNRPQVRPL  
SPGESGAWDIPIGGIMPGRYNQEVGQTIPVFAFLGAMVVLAFFVUQINKAKSRPKRRKPRVKRQPL  
MQQVHPPKTPSV

FIG. 7

14/16

	10	20	30	40	50
BPN' SAVINASE S2HSBT	AQ S V P Y G V S Q - I K A P A L H S Q G Y T G S N V K V A V I D S G I D S S H P D L K - V A G G A 48 A Q S V P W G I S R - V Q A P A A H N R G L T G S G V K V A V L D T G I - S T H P D L N - I R G G A 47 - R A I P R Q V A Q T L Q A D V L W Q M G Y T G A N V R V A V F D T G L S E K H P H F K N V K E R T 49				
	60	70	80	90	100
BPN' SAVINASE S2HSBT	S M V P S E T N P F Q D N N S H G T H V A G T V A A L N N S I G V L G V A P S A S L Y A V K V L G A 98 S F V P G E P S T - Q D G N G H G T H V A G T I A A L N N S I G V L G V A P S A E L Y A V K V L G A 96 N W - - T N E R T L D D G L G H G T F V A G V I A S M R E C Q G F - - - A P D A E L H I F R V F T N 94				
	110	120	130	140	150
BPN' SAVINASE S2HSBT	D G S G O Y S W I I N G I E W A I A N N M D V I N M S L G G P S - G S A A L K A A V D K A V A S G V 147 S G S G S V S S I A Q G L E W A G N N G M H V A N L S L G S P S - P S A T L E Q A V N S A T S R G V 145 N Q V S Y T S W F L D A F N Y A I L K K I D V L N L S I G G P D F M D H P F V D K V W E L T A N N V 144				
	160	170	180	190	200
BPN' SAVINASE S2HSBT	V V V A A G N E G T S G S S S T V G Y P G K Y P S V I A V G A V D S S N Q R A S F S S V G P E L - 197 L V V A A S G N S G A - - - G S I S Y P A R Y A N A M A V G A T D Q N N N R A S F S Q Y G A G L - 191 I M V S A I G N D G P - - - L Y G T L N N P A D Q M D V I G V G G I D F E D N I A R F S S R G M T T W 192				
	210	220	230	240	250
BPN' SAVINASE S2HSBT	- - - - - D V M A P G V S I Q S T L P G N K Y G A Y N G T S M A S P H V A G A A A L I L 235 - - - - - D I V A P G V N V Q S T Y P G S T Y A S L N G T S M A T P H V A G A A A L V K 229 E L P G G Y G R M K P D I V T Y G A G V R G S G V K G G C R A L S G T S V A S P V V A G A V T L V 242				
	260	270	280	290	
BPN' SAVINASE S2HSBT	S K H P N W T N T Q - - - V R S S L E N T T K L G D S F Y Y G K G L I N V Q A A A Q 275 Q K N P S W S N V Q - - - I R N H L K N T A T S L G S T N L Y G S G L V N A E A A T R 269 S T V Q K R E L V N P A S M K Q A L I A S A R R L P G V N M F E Q G - - - H G K L 280				

FIG. 8

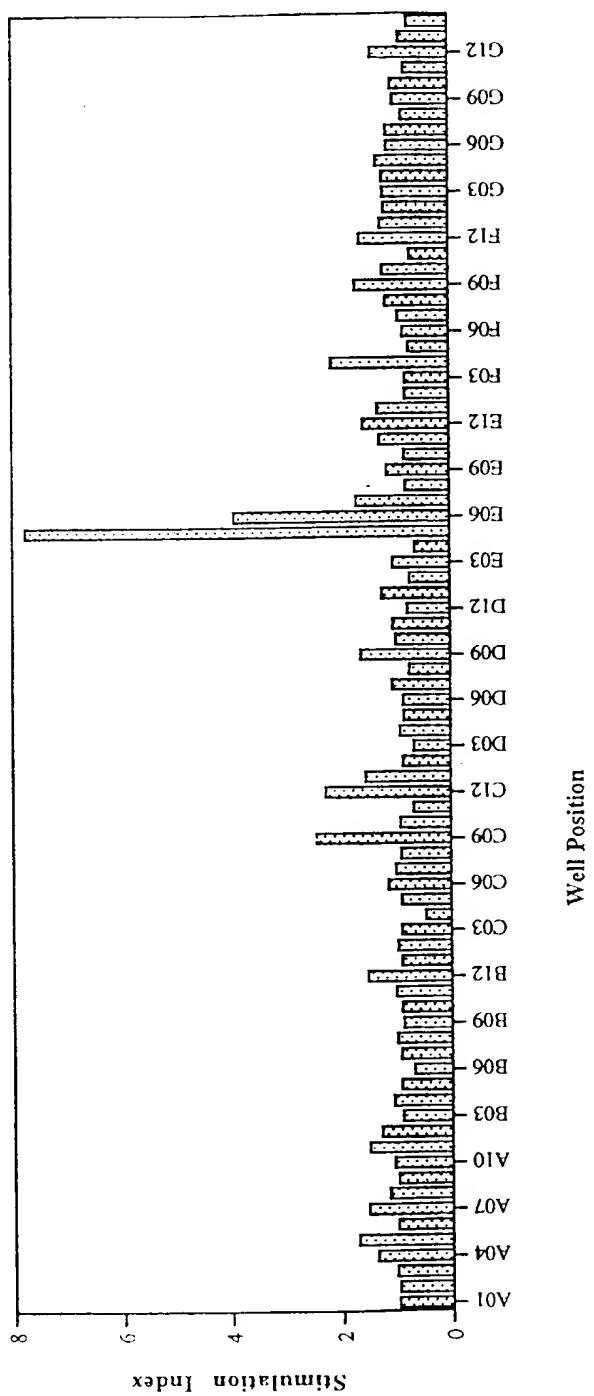


FIG. 9

16/16

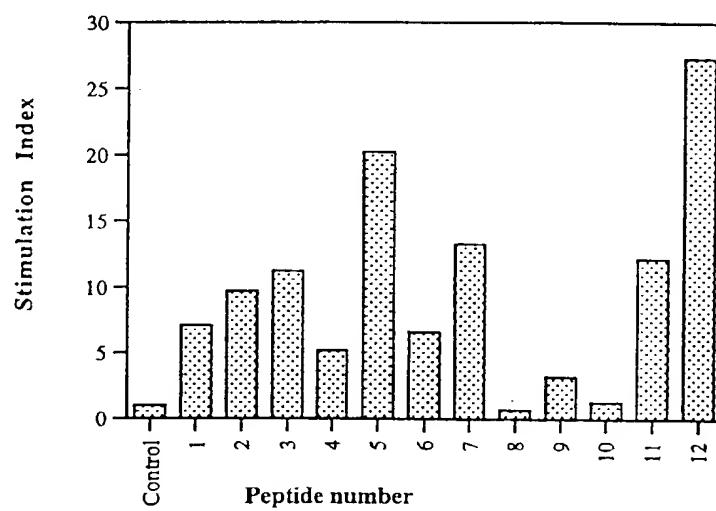


FIG. 10